Characterization and Investigation of BBa_I0500 (Inducible pBad/araC promoter)

* For proper nomenclature, we will refer to this BioBrick as *araC-P$_{BAD}$*

* We use the term sensing range instead of input dynamic range (IDR) because the latter specifically refers to a ratio between values rather than the values themselves.

Introduction

This characterization on BBa_I0500 *araC-P$_{BAD}$* promoter (or *araBp* promoter) is meant to complement our previous characterization submitted in 2014. We discovered a few issues in reproducibility and interpretation. Essentially, our result from 2014 did not fully capture the arabinose sensing range. The all-or-none response reported on pSB3K3 was also an artifact, where dead cells from an old plate interfered with fluorescent readings. To provide characterization with quality within limited time, we decided to go for a deeper characterization instead of measuring multiple properties (e.g. induction time). Our results demonstrated that this promoter 1) has a graded induction response, 2) has different sensing ranges on plasmids with different copy numbers, and 3) does not display all-or-none behavior on a single cell level when expressed from a low copy plasmid.

Alongside the characterization, we also report our misuse of this promoter and issues in its sequence annotation. We hereby present our data and caution users in choosing this promoter in future applications.

Results:

1. The induction by arabinose on population level expression is a graded response

The transfer function of BBa_I0500 was produced using 24 arabinose concentrations from 16mM to 0.00381μm, using log-2 spaced concentrations. Molar concentration was chosen instead of percentage (weight by volume) because it was more commonly used among experimental biologists. The molar concentration to percentage conversion was provided instead for reference.
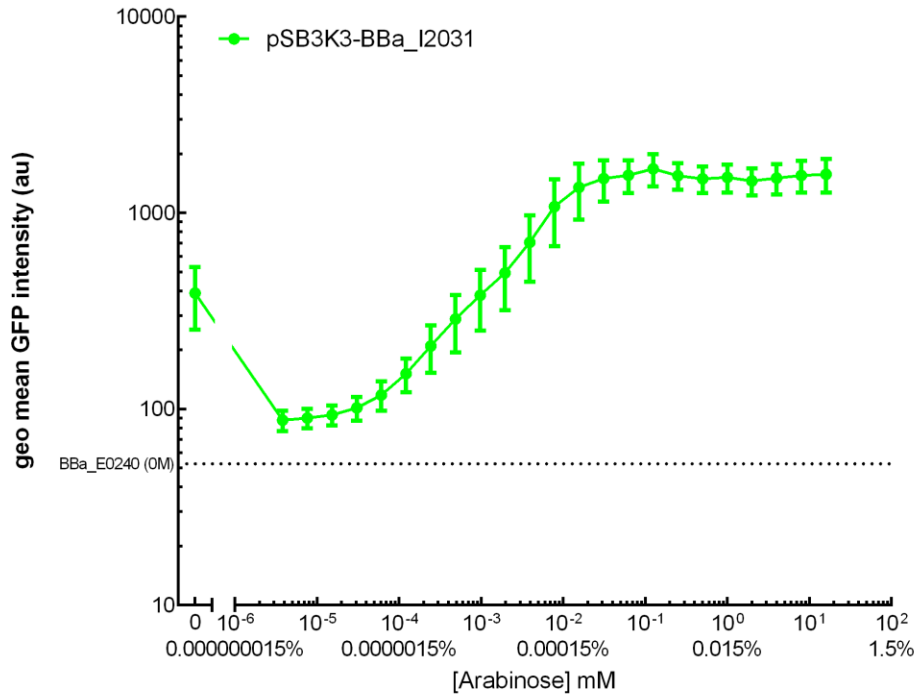

Figure 1. Transfer function for BBa_I2031 promoter on plasmid pSB3K3. The dashed line represents cells' auto fluorescence, measured using DH10B cells harboring pSB3K3-BBa_E0240. Error bars represent SEM of 3 independent experiments on 3 different days.

Our result (Figure 1) shows that the induction for BBa_I0500 is graded and roughly spans from 10^{-4} mM (or 0.0001mM) to 10^{-2} mM (or 0.01mM), and the production of GFP saturates beyond these concentrations. The GFP intensity at 0mM is aberrantly higher than those with low concentrations of arabinose. We could not explain this weird data point but we could be certain that it was not due to mistakes in experiment execution, because the same batch of medium was used in characterization of BBa_I2031 on pSB1K3 plasmid (please refer to below), and when the same medium with 0M arabinose was used, no significant induction was observed.

Our result agrees with neither Cambridge 2011's nor Groningen 2011's results (Figure 2):

- In comparison to result by Cambridge 2011:

Maximum expression of GFP can also be seen at 0.01mM arabinose. Yet, expression is not fully off at 0.001mM of arabinose. The claimed call all-or-none response was not observed as well – the fluorescence intensity on a population level can be tuned incrementally.

In the reference that Cambridge 2011 cited (Khlebnikov et al., 2001), the "all-or-none" response refers to an on-off switching behavior of P_{BAD} on a single cell level but not on a population level. We believe the statement "response to arabinose is highly variable between different strains of *E coli*" was an unintentional misconstrue of the fact that, in BW27783, constitutive expression of AraE would convert the single cell all-or-none response into a graded response, and that should be irrelevant to the induction response on a population level. The issue of all-or-none response on a single cell level will be revisited in a separate part of this report. Given this, we believe characterization using a different strain here is not the source of discrepancy in our results.

- In comparison to result by Groningen 2011:

Groningen 2011 reported an arabinose sensing range from 0.05% to 1% arabinose (corresponding to 3.33 mM and 66.7 mM respectively). That reported range lies in our induction saturated range.

In our effort to reconcile previous characterization data, it appeared that we have created more disagreeing results instead.

Note: We also noticed that there are other versions of *araC-P_{BAD}* in the Registry, notably BBa_K1067007 (sensing range uncharacterized) and BBa_K808000 (sensing range from 0.01% to 1% arabinose). However, their sequences were not identical to BBa_I0500 and were thus omitted from this comparison.
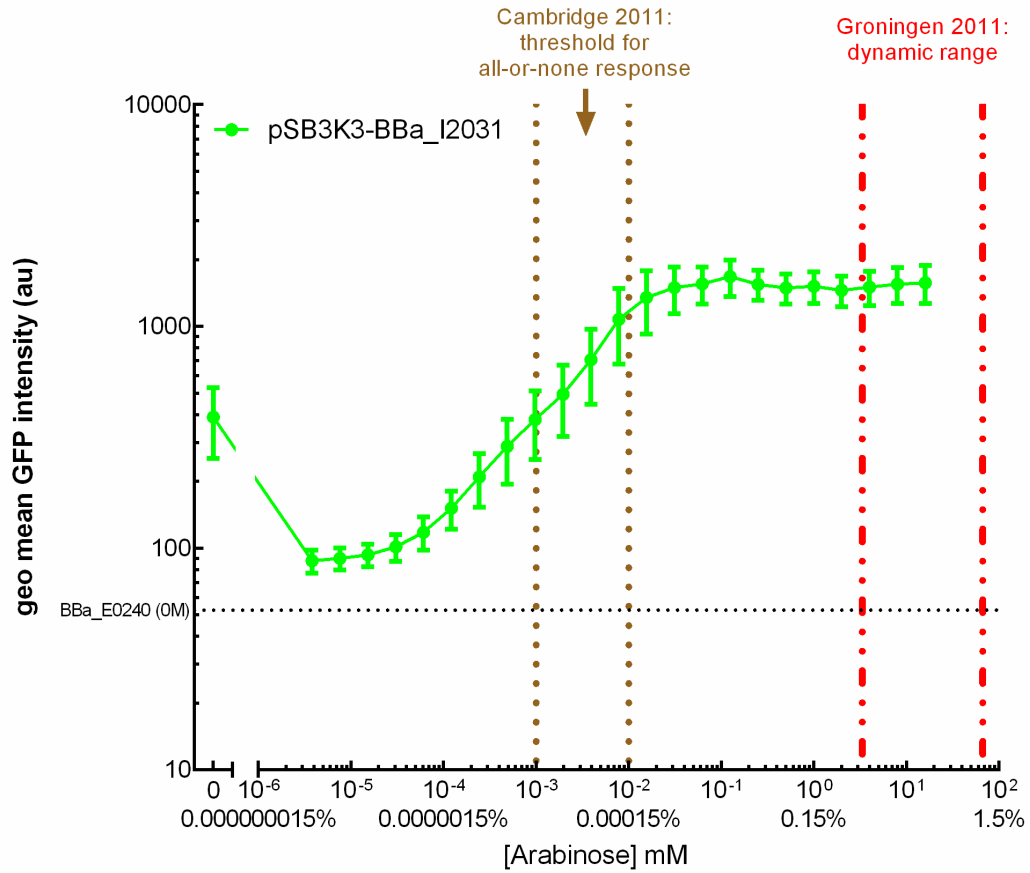
Figure 2. Comparison of our result to those by Cambridge 2011 and Groningen 2011. Cambridge 2011 reported a threshold of all-or-none between 0.001mM and 0.01mM of arabinose and Groningen 2011 reported a sensing range from 0.05% to 1% arabinose. Our result did not agree with either of theirs.

2. BBa_I0500 on a high copy plasmid is sensitive to higher ranges of arabinose

Groningen 2011 team used the high copy pSB1C3 plasmid for their characterization instead of the conventional low copy pSB3K3 plasmid. It was thus speculated that the discrepancy between the results from Groningen 2011 and Cambridge 2011 was due to use of plasmids with different copy numbers.

The BBa_I2031 construct was migrated from pSB3K3 to pSB1K3 and was characterized in its then high copy context.
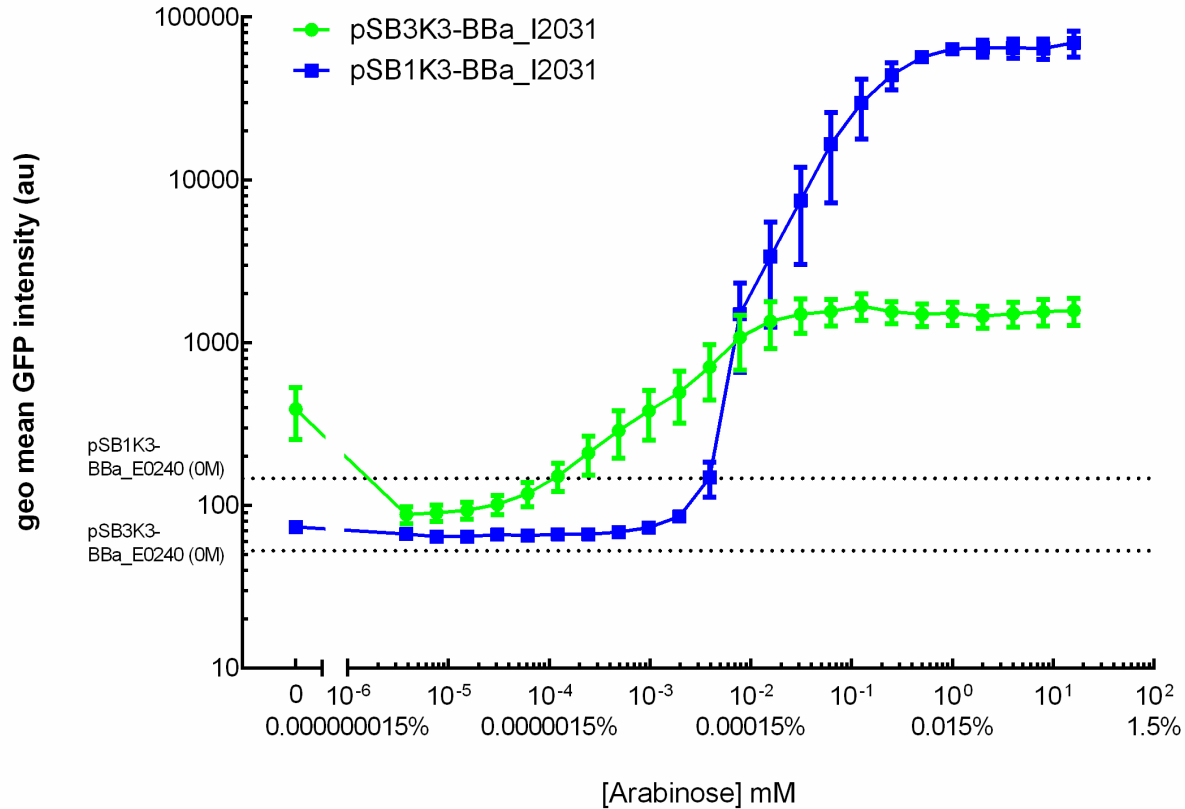

Figure 3. Transfer functions for BBa_I2031 on plasmid pSB3K3 and pSB1K3. The dashed lines represent cells' auto fluorescence, measured using DH10B cells harboring pSB3K3-BBa_E0240 or pSB1K3-BBa_E0240. Error bars represent SEM of 3 independent experiments on 3 different days.

Without surprises, stronger fluorescence can be observed from fully induced BBa_I0500 on pSB1K3 (Figure 3). Nonetheless, the two transfer functions showed some interesting differences:

1. On pSB3K3, BBa_I0500 is responsive to 10^{-4} - 10^{-2} mM arabinose, whereas on pSB1K3, it senses arabinose from roughly 10^{-3} to 1mM (see Figure S1).
2. BBa_I2031 appeared to be giving less fluorescence in low arabinose concentrations when placed on the high copy pSB1K3 plasmid than on low copy pSB3K3 plasmid, and the value is even lower than the auto fluorescence observed from the negative control (DH10B / pSB1K3-BBa_E0240). It might be interpreted that the promoter is less leaky when placed on a high copy plasmid.

3. BBa_I0500 on a low copy plasmid does not display all-or-none behavior on a single cell level

The all or none behavior of P_{BAD} on a single cell level has been reported in a number of literatures (Fritz et al., 2014; Khlebnikov et al., 2000; Khlebnikov et al., 2001; Siegele and Hu, 1997). It describes that an increase in arabinose concentration does not result in an increase in promoter / gene activity per se, but rather, increases the proportion of population that are fully induced (Figure 4). According to Khelbnikov et al., the all-or-none behavior observed is due to the autocatalytic behavior of the AraE transporter, creating a positive feedback mechanism on a single cell level. Driving expression of the AraE by a constitutive promoter abrogates the feedback and transforms the all-or-none behavior into homogenous expression of P_{BAD}. (Khlebnikov et al., 2001)
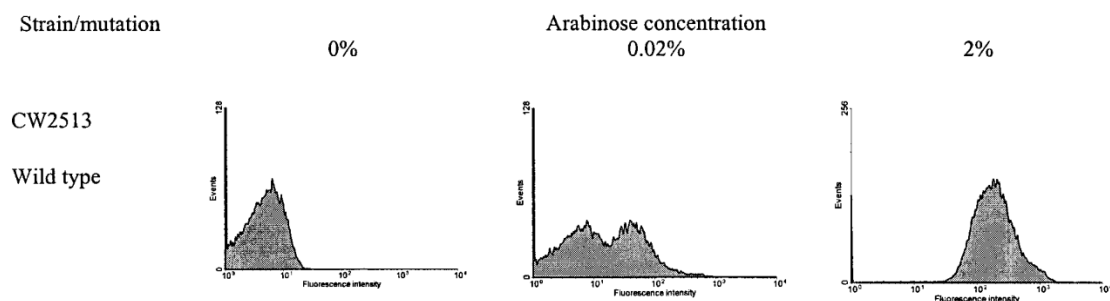

Figure 4. Figure from literature showing all-or-none behavior of P_{BAD} on a single cell level. As shown in previous studies, arabinose induction at a sub-saturating concentration leads to heterogeneously activated population. Figure was directly cropped from figure 3 in the paper "Regulatable Arabinose-Inducible Gene Expression System with Consistent Control in All Cells of a Culture" (Khlebnikov et al., 2000).

Nonetheless, as revealed by our histograms for intermediate level of induction, the situation described above is not always true. The all-or-none behavior applied only when BBa_I0500 is placed on a high copy plasmid, where induced cells were mostly distributed among two bins of fluorescence (Figure 5). Yet, when the low copy pSB3K3 plasmid was used, the all-or-none behavior no longer holds and the populations remain homogenous along the arabinose concentration gradient.

Going back to the literature, we confirmed that all of them were studying P_{BAD} on plasmids with high copy origins of replication (Table 1). Thus, we believe that the homogenous expression of P_{BAD} promoter from a low copy plasmid has been a long overlooked issue.

Reference	Plasmid harboring P_{BAD}-reporter	Plasmid lineage	Traced origin of replication
Khlebnikov et al., 2000 Khlebnikov et al., 2001	pCSAK50	pBAD24 → pTC40 → pCSAK50	High copy pBR322 origin *
Siegele and Hu, 1997	pDS439	pBAD18 → pDS439	High copy pBR322 origin *
Fritz et al., 2014	pBAD24-GFP	pBAD24 → pBAD24-GFP	High copy pBR322 origin *

* The paper by Guzman et al. described the origins as pBR origins (Guzman et al., 1995). However, they lack the *rop* gene that maintains low copy number. Thus they are high copy origins (Cronan, 2006). The origin of pBAD24 in ATCC (ATCC® 87399™) was also documented to have the pMB1 origin, which should be the same origin in pSB1C3.

Table 1. Replication origins used in previous studies of P_{BAD} promoter. As shown, all were derived from the pBAD plasmid series and are thus high copy number plasmids.
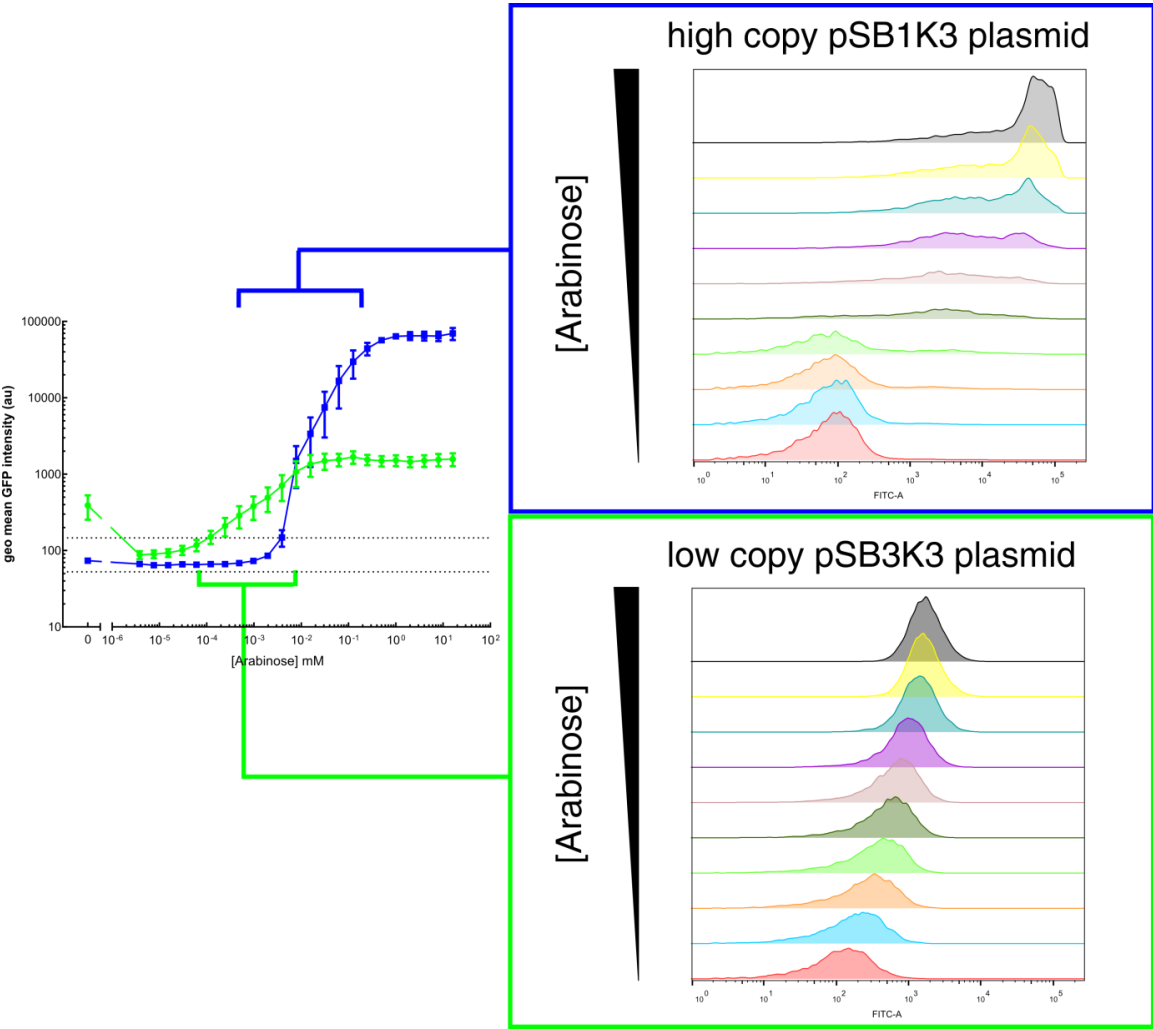

Figure 5. Histogram plots for sensing ranges of BBa_I0500 on high and low copy plasmid. Concentrations of arabinose for high copy pSB1K3 plasmid: 0.488µM – 0.25mM. For low copy pSB3K3 plasmid: 0.0610µM – 0.03125mM. Only 1 set of experiment result from 3 replicates is presented.

4. Our misuse of BBa_I0500 and its part documentation issue

BBa_I0500 has been a popular part with 329 uses to date (as of 4th Sep 2015), and has functioned as expected in many cases. Unfortunately, we have had unpleasant experiences with BBa_I0500 because we were unaware of one of its properties, which was not documented on its Registry Page.

In the year 2014, our iGEM team was trying to reproduce previous riboregulators results from different iGEM teams, as well as the very first versions of riboregulators, crR12 and taR12 (Isaacs et al., 2004). Isaacs et al. has used a non-BioBricked version of P_{BAD} to control expression of taR12. To follow the setup as close as possible, we employed BBa_I0500 to drive expression of key1 (BBa_J01008) in presence of constitutively expressed lock1-RBS-GFP (BBa_K1379020). However, our results did not show a significant upregulation of gene expression upon induction with arabinose.

A year later, now in 2015, we tried to troubleshoot and we identified the design of BBa_I0500 as a possible cause of failure. Issacs et al. demonstrated in their paper that a precise 5' end for the taRNA is indispensable and additional nucleotides can jeopardize its ability to open up the crRNA hairpin structure (Isaacs et al., 2004). BBa_I0500 contains 19 additional nucleotides after its native transcription start site, and thus would add them to the 5' end of taRNA and should render it functionless.

In addition, the extra nucleotides also invalidated our attempt to characterize the Relative Promoter Unit (RPU) for BBa_I0500, because one of the prerequisites for promoter-BBa_E0240 to be made comparable to the standard promoter BBa_I20260 is that they drive transcription of mRNA with identical sequences (Kelly et al., 2009). With the extra nucleotides, BBa_I0500 is not a valid promoter for standard promoter measurement.

The lack of sequence feature information might also have impacted the design of BBa_I0500's derivative BBa_I13453, which is a truncated form of BBa_I0500. It was created with the goal of decoupling the P_{BAD} promoter from the *araC* operon. In its experience page, iGEM 2012 Michigan analyzed the sequence and reported that the O2 binding site is missing. Theoretically, AraC mediated repression in absence of arabinose should not happen. Interestingly though, when put into MG1655, where an endogenous copy of AraC is present, the promoter appears to be inducible (British_Columbia 2009, K.U.Leuven 2011, SDU-Denmark 2012). However, if the information of the O2 binding site was available earlier, we believe the part could have been more carefully designed to fully utilize the regulation by AraC.

In summary, our experience of failure is due to the incognizance of sequence features in BBa_I0500, leading to its misuse, and we believe this can be improved with proper annotation. Yet due to access authority settings, we could not add back the sequence features to BBa_I0500. In view of this, we strongly recommend future users to cross check the sequence features of BBa_I0500 with that of BBa_K1067007, which has most, if not all, of the important sequence features annotated.

Materials and Methods

Strain and Medium

All measurements were done in *E. coli* DH10B cells growing in M9 medium supplemented with 0.2% casamino acids, 4% glycerol and 50μg/mL kanamycin (hereto referred as M9s, please refer to our protocol page in wiki for the recipe) with various concentrations of L-(+)-Arabinose (Sigma #A3256). Graded concentrations of M9s medium with arabinose were prepared by serial dilution from 16mM M9s+arabinose in a 2-fold manner.

Measurement

For each independent experiment, a single colony of each strain was picked and inoculated in M9s until a turbid culture was obtained. 50 μL starter cultures were then inoculated in 500 μL of different concentrations of fresh M9s+arabinose for an overnight period at 37°C with 200rpm shaking, in a 96 well deep plate (Sigma #CLS3960) sealed with a breathable tape (Sigma #CLS3345). This step is to minimize the lag time by arabinose induction so that cells do not need to re-adjust to new arabinose concentrations, and hopefully that they would be attaining induction equilibrium when reaching log phase. The next day, the culture was diluted 20-fold into a new plate with fresh M9s+arabinose medium, each into their corresponding and identical arabinose concentration. The new plate was then sealed and inoculated in an identical manner as the overnight culture. OD_{600} was monitored by removing a 200 μL aliquot from the culture and then reading in a plate reader. When most culture in the wells reached OD_{600} from 0.25 to 0.5, 300 μL of cells would then be fixed by diluting into 300 μL of 1% paraformaldehyde in 1X PBS.

We used the BD FACSAria IIIu Cell Sorter for all measurements, with the following settings: FSC 300V, SSC 300V, FITC 500V. Samples were run at 200-400 events per second on a flow rate of 10 and 10000 events were collected for each sample. Noise was distinguished from background fluorescence using DH10B / pSB3K3-BBa_E0240 and was gated away. The remaining population geometric mean fluorescence values and the histograms were then exported for analysis.
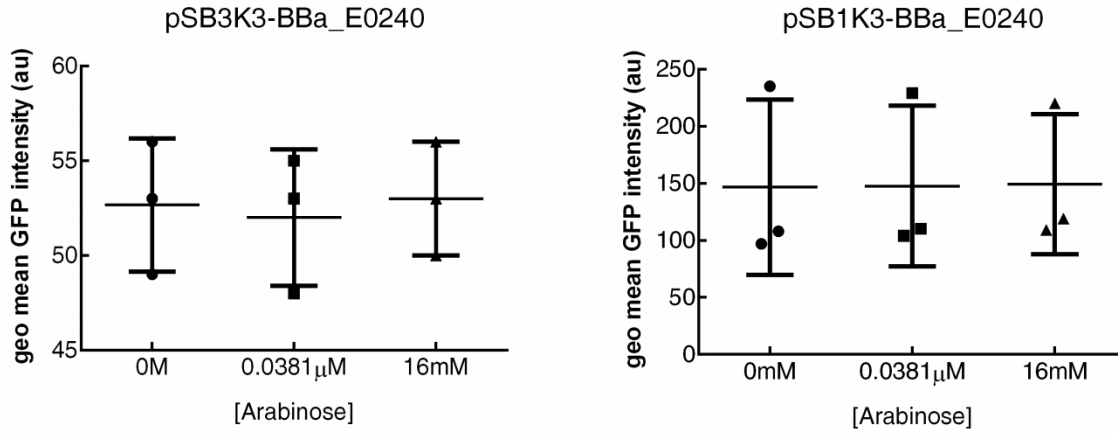
Supplementary information

Figure S1. Background fluorescence measurement using BBa_E0240 on pSB3K3 and pSB1K3. DH10B cells harboring pSB3K3-BBa_E0240 or pSB1K3-BBa_E0240 were grown in 3 concentrations of arabinose in M9s medium and measured as described. This shows that the background fluorescence is not affected by arabinose concentration and therefore we took the fluorescence value at 0M arabinose as the background fluorescence. Each point represents a data point collected in an independent experiment. Error bar represents SEM.

Reference List

Cronan, J.E. (2006). A family of arabinose-inducible Escherichia coli expression vectors having pBR322 copy control. Plasmid *55,* 152-157.

Fritz, G., Megerle, J.A., Westermayer, S.A., Brick, D., Heermann, R., Jung, K., Rädler, J.O., and Gerland, U. (2014). Single Cell Kinetics of Phenotypic Switching in the Arabinose Utilization System of *E. coli*. PLoS ONE *9,* e89532.

Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. Journal of Bacteriology *177,* 4121-4130.

Isaacs, F.J., Dwyer, D.J., Ding, C., Pervouchine, D.D., Cantor, C.R., and Collins, J.J. (2004). Engineered riboregulators enable post-transcriptional control of gene expression. Nat Biotech *22,* 841-847.

Kelly, J., Rubin, A., Davis, J., Ajo-Franklin, C., Cumbers, J., Czar, M., de Mora, K., Glieberman, A., Monie, D., and Endy, D. (2009). Measuring the activity of BioBrick promoters using an in vivo reference standard. Journal of Biological Engineering *3,* 4.

Khlebnikov, A., Datsenko, K.A., Skaug, T., Wanner, B.L., and Keasling, J.D. (2001). Homogeneous expression of the PBAD promoter in Escherichia coli by constitutive expression of the low-affinity high-capacity AraE transporter. Microbiology *147,* 3241-3247.

Khlebnikov, A., Risa, Ø, Skaug, T., Carrier, T.A., and Keasling, J.D. (2000). Regulatable Arabinose-Inducible Gene Expression System with Consistent Control in All Cells of a Culture. Journal of Bacteriology *182,* 7029-7034.

Siegele, D.A., and Hu, J.C. (1997). Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations. Proceedings of the National Academy of Sciences *94,* 8168-8172.